Guatemala - Education

Report generated on: June 26, 2018

Visit our data catalog at: https://data.mcc.gov/evaluations/index.php

Overview

Identification

COUNTRY
Guatemala

EVALUATION TITLE
Education

TRANSLATED TITLE
El Programa Umbral en Guatemala: Informe de Diseño de Evaluación

EVALUATION TYPE
Independent Evaluation

ID NUMBER
DDI-MCC-GTM-MPR-EDU-2018-v01

Version

VERSION DESCRIPTION
Anonymized dataset for public distribution

Overview

ABSTRACT
Activity 1:

We will estimate Activity 1's causal impacts using a randomized controlled trial. The Ministry of Education (MINEDUC) grouped all eligible schools within the five study departments into school districts-clusters of neighboring schools with eight schools each, on average. We randomly assigned 80 percent of those districts to a treatment group, in which schools would be invited to participate in Activity 1. We assigned the remaining 20 percent of districts to a control group, in which the implementer would not intervene.

We will gather baseline and endline data from a randomly selected sample of schools from the treatment and control groups. We will estimate Activity 1's causal effects by comparing outcomes from the two groups at endline and controlling for chance differences observed at baseline. We will report two types of estimates: intent-to-treat estimates, which reflect differences between treatment and control schools regardless of their participation in treatment, and local average treatment effects, which are adjusted to account for participation rates in both groups. We describe our sample selection, randomization, and data collection methods elsewhere in sampling and data collection methods sections.

We will describe the implementation of the activity and identify its facilitators and barriers via an implementation study. We describe our sample selection and data collection methods in the sampling and data collection methods sections.

Activity 3:

We will conduct a trend analysis to understand changes in key outcomes related to this activity. This approach will not show the activity's causal impacts, but we will combine the trend analysis with an analysis of related concurrent events to see the potential relationship between the activity and outcomes of interest.

We will also conduct a qualitative analysis, relying on focus groups, interviews, and implementation documentation. We will use the drivers-of-change (DOC) framework and political economy methods to structure the qualitative component. The DOC framework is an analytic framework for applying political economy analysis, which enables evaluators to systematically assess how project design and implementation decisions addressed contextual factors that may affect whether project goals are met. Specifically, the DOC framework calls for an analysis of the relationship between structural features (such as the structure and history of education in Guatemala and social and demographic trends), institutions (such as the legal framework, government policies, and informal rules that affect behavior), and agents (such as organizations and people who

influence and participate in lower-secondary education).

Note:

Under Evaluation Methodology, we selected Randomization, which is the method for the impact evaluation of Activity 1. We will also conduct implementation evaluations of Activities 1 and 3 and political economy analyses of Activity 3.

Under data type, we selected Sample Survey Data, which is the main method for the impact evaluation of Activity 1. For both activities, we will also use administrative data, monitoring and evaluation data from the implementer, and qualitative data.

EVALUATION METHODOLOGY
Randomization

UNITS OF ANALYSIS
Unit of Analysis

Activity 1: The unit of randomization is the school district, a group of schools that includes eight schools on average. We will conduct analyses at the school, teacher, and student levels.

Activity 3: The analysis will be national in scope. If possible (depending on data availability and relevance), we may focus some of the analysis on the five departments that are the focus of Activity 1 (Alta Verapaz, Chiquimula, Jalapa, Sololá, and Sacatepéquez).

Note:

Kind of Data

Activity 1:

Survey data (teacher and school director), classroom observation data, student test score data, administrative data, monitoring and evaluation data, focus group data, and interview data.

Activity 3:

Administrative data, monitoring and evaluation data, and interview data.

KIND OF DATA
Sample survey data [ssd]

TOPICS

Topic	Vocabulary	URI
Education	MCC Sector	

KEYWORDS
Lower-secondary education, Randomized controlled trial, Impact evaluation, Performance evaluation, Political economy analysis, School networks, Parent councils, Pedagogic support, Teacher diagnostics, Management information systems, Student learning, Student retention and promotion, Gender and social inclusion, Éxito Escolar, Strengthening of Institutional and Planning Capacity

Coverage

GEOGRAPHIC COVERAGE
Activity 1: Departments of Alta Verapaz, Chiquimula, Jalapa, Sololá, and Sacatepéquez. All lower-secondary schools in these departments, excluding private and municipal schools, are eligible for Activity 1 and are included in the randomized controlled trial. We will gather survey and classroom observation data from a random sample of 332 schools in these departments, and we will conduct focus groups and interviews with stakeholders from a subset of the schools in the data collection sample.

Activity 3: National

UNIVERSE
Activity 1, quantitative impact evaluation

The universe of eligible schools includes all public lower-secondary schools in the five departments of interest (Alta Verapaz, Chiquimula, Jalapa, Sacatepéquez, and Sololá), including cooperative schools but excluding municipal schools. MINEDUC grouped these 805 schools into 103 districts. Each district includes 8 geographically proximate schools, on average.

The implementer, FHI360, noted the need to include roughly 80 percent of teachers in the treatment group in order to reach its target number of teachers for the teachers' professional development subactivity. To accommodate this, we assigned 80 percent of districts (82 districts with 631 schools) to the treatment group. The treatment group includes 80.5 percent of the teachers in study schools (3,490 of 4,337 teachers). The control group includes 21 districts with 174 schools. We conducted random assignment via a public lottery in Guatemala.

For survey data collection and classroom observations, our sample includes all schools in districts assigned to the control group (174 schools) and one-quarter of the schools in districts assigned to the treatment group (158 schools). We drew the sample in such a way as to include at least one school from every district in the study. At each sampled school, our data collection partner, Espirálica, will survey the school director and will survey and observe one randomly selected 7th-grade teacher from each of three subjects: language and communication, math, and natural sciences. In small schools, one teacher may cover more than one of the subjects. Enumerators will select teachers from a roster that they will prepare with the school director when they visit the school for data collection.

MINEDUC will gather baseline test-score data from all 7th-grade students and from one section of 9th-grade students in the 332 schools selected for data collection. At endline, we will collect data from school directors and 9th-grade teachers, and we expect MINEDUC to collect student test-score data from all 9th-grade students at the 332 schools. MINEDUC may also collect student test-score data from 7th-grade students at endline.

Activity 1, qualitative evaluation of implementation

For the implementation evaluation of the Éxito Escolar activities, we will draw a purposeful sample of school directors, teachers, students, and parents from the treatment and control groups to participate in focus groups (and interviews, in the case of school directors). For teachers and students, we will conduct up to 15 focus groups with treatment schools and 6 with control schools at baseline and endline. For parents, we will conduct up to 9 focus groups each at baseline and endline (treatment schools only). We will also interview up to 15 treatment school directors and 6 control school directors at baseline and endline. In addition, we will interview representatives of MINEDUC and FHI360 (the implementer).

Activity 3:

We will interview key informants at the national level, in two participating regions, and possibly in subregions, depending on the decision-making locus.

Producers and Sponsors

PRIMARY INVESTIGATOR(S)

Name	Affiliation
Mathematica Policy Research	

FUNDING

Name	Abbreviation	Role
Millennium Challenge Corporation	MCC	

Metadata Production

METADATA PRODUCED BY

Name	Abbreviation	Affiliation	Role
Mathematica Policy Research	MPR		Review of Metadata

DATE OF METADATA PRODUCTION

2018-06-22

DDI DOCUMENT VERSION
Evaluation Design Report, Version 1

DDI DOCUMENT ID
DDI-MCC-GTM-MPR-EDU-2018-v01

MCC Compact and Program

COMPACT OR THRESHOLD
Guatemala Threshold

PROGRAM
Overall: The objective of the Guatemala Threshold Education Project is to support the government of Guatemala's ongoing institutional reforms, as defined in the Proposal for the Transformation of Secondary Education (Asturias de Barrios 2014) and the Ruta Crítica (MINEDUC 2016). The goal of these reforms is to provide youth in Guatemala with a high quality secondary education that prepares them for success in the labor market. The project consists of three complementary activities. This Nesstar entry provides information on the evaluations of Activities 1 and 3, described below. FHI360 will implement the project. Activity 1 (Éxito Escolar): Activity 1 is the Éxito Escolar activity. It includes four components, all provided by FHI360, that are intended to improve the quality of education at the lower-secondary level. The first component is an intensive, 20-month professional development program for lower-secondary school teachers and principals, including in-person and virtual elements. The second component is pedagogic support for participating schools and the formation of communities of practice, through which teachers can engage with teachers at neighboring schools. The third component is school networks, which will link primary schools to a neighboring lower-secondary school and link lower-secondary schools to a neighboring upper-secondary school. One purpose of these networks is to create a feeder system that improves student retention and the transition from primary to lower-secondary education. The fourth component is the formation of parent councils at some participating schools, which will train parents on how to support their schools and help prevent dropout. Activity 3 (Strengthening of Institutional Capacity and Planning): Activity 3 is the Strengthening of Institutional Capacity and Planning (IPC, for its name in Spanish) activity. Its objective is to strengthen the institutional capacity of MINEDUC to improve its planning and budgeting functions, with the ultimate goal of enabling it to provide equitable and high quality secondary education (MCC 2016). The activity includes four main subactivities. The first is to identify the minimum inputs required for a high quality education and the budget needed to ensure the minimum inputs in the secondary education system. The second is to strengthen management information systems, including the creation of a virtual bank of education resources available to school directors, teachers, parents, and other community members. The third is to advance the institutionalization of a competitive teacher-selection process (including a diagnostic test). And the fourth is to develop a geographic analysis of the supply and demand of secondary education as a factor in estimating needed resources for infrastructure, teacher assignments, and materials and in planning and budgeting for a high quality education. Note: We list the MCC sector as Education, but Activity 3 could fall under education or Capacity Building and Institutional Development. Asturias de Barrios, Linda. “Propuesta de transformación del nivel de educación media: versión revisada y validada. Documento para la discusión.” Guatemala City, Guatemala: USAID, January 2014. MINEDUC. “Ruta Crítica: Avances y Retos. Prioridades del Plan de Implementación Estratégica de Educación 2012-2016.” Guatemala: Ministerio de Educación, July 2014. Millennium Challenge Corporation. “MCC-16_RFQ-0131 (Guatemala Education Project).” Washington, DC: May 24, 2016.

MCC SECTOR
Education (Edu)

PROGRAM LOGIC
Activity 1: The program logic of Activity 1 assumes that the interventions will enhance the quality of education in the short term (leading to improved student learning) and increase retention and promotion in the medium term (MCC 2016). The program logic further assumes that improvements in students' outcomes should produce graduates who are better prepared for the workforce over the long term. This assumption links the activity to the main project objective of improving the education of Guatemalan youth to prepare them for job success. Activity 3: The program logic of Activity 3 envisions institutional-strengthening efforts that will integrate the various secondary education modalities and improve equity and results. This will be accomplished by developing and implementing policies, systems, and tools to recruit teachers and to allocate material and financial resources effectively and equitably. Millennium Challenge Corporation. “MCC-16_RFQ-0131 (Guatemala Education Project).” Washington, DC: May 24, 2016.

PROGRAM PARTICIPANTS
Activity 1: The main participants in Activity 1 are (1) lower-secondary school directors and teachers who participate in activities such as professional development, pedagogic support, or communities of practice and (2) their students, who may benefit from improved school management and teaching. Additional participants include parents who take part in parent

councils and primary school directors or teachers who engage in school network activities (we will learn about the experiences of these groups through our qualitative research, but our impact evaluation will not include estimating impacts on parents or primary school staff). Activity 3: Because Activity 3 will engage MINEDUC staff in improving the planning and management of lower-secondary education, it will directly affect a wide range of participants. It will involve training 50 directors and technicians in MINEDUC's central headquarters, as well as department directors who plan and manage lower-secondary education. The activity will include at least two education budget workshops, with about 1,000 participants per workshop. Interested stakeholders from civil society and the private sector will be permitted to take part in some of the institutional capacity-building workshops. Given that MINEDUC is responsible for lower-secondary education throughout Guatemala, Activity 3 will be national in scope. Potential indirect beneficiaries include school directors, teachers, parents, and community members who access the virtual bank of education resources to be established under this activity. These groups would benefit because as the government allocates more resources to the schools, the additional resources could increase teacher pay, the number of qualified teachers, and the provision of professional development. However, we will not measure or quantify these benefits in the evaluation.

Sampling

Study Population

Activity 1, quantitative impact evaluation The universe of eligible schools includes all public lower-secondary schools in the five departments of interest (Alta Verapaz, Chiquimula, Jalapa, Sacatepéquez, and Sololá), including cooperative schools but excluding municipal schools. MINEDUC grouped these 805 schools into 103 districts. Each district includes 8 geographically proximate schools, on average. The implementer, FHI360, noted the need to include roughly 80 percent of teachers in the treatment group in order to reach its target number of teachers for the teachers' professional development subactivity. To accommodate this, we assigned 80 percent of districts (82 districts with 631 schools) to the treatment group. The treatment group includes 80.5 percent of the teachers in study schools (3,490 of 4,337 teachers). The control group includes 21 districts with 174 schools. We conducted random assignment via a public lottery in Guatemala. For survey data collection and classroom observations, our sample includes all schools in districts assigned to the control group (174 schools) and one-quarter of the schools in districts assigned to the treatment group (158 schools). We drew the sample in such a way as to include at least one school from every district in the study. At each sampled school, our data collection partner, Espirálica, will survey the school director and will survey and observe one randomly selected 7th-grade teacher from each of three subjects: language and communication, math, and natural sciences. In small schools, one teacher may cover more than one of the subjects. Enumerators will select teachers from a roster that they will prepare with the school director when they visit the school for data collection. MINEDUC will gather baseline test-score data from all 7th-grade students and from one section of 9th-grade students in the 332 schools selected for data collection. At endline, we will collect data from school directors and 9th-grade teachers, and we expect MINEDUC to collect student test-score data from all 9th-grade students at the 332 schools. MINEDUC may also collect student test-score data from 7th-grade students at endline. Activity 1, qualitative evaluation of implementation For the implementation evaluation of the Éxito Escolar activities, we will draw a purposeful sample of school directors, teachers, students, and parents from the treatment and control groups to participate in focus groups (and interviews, in the case of school directors). For teachers and students, we will conduct up to 15 focus groups with treatment schools and 6 with control schools at baseline and endline. For parents, we will conduct up to 9 focus groups each at baseline and endline (treatment schools only). We will also interview up to 15 treatment school directors and 6 control school directors at baseline and endline. In addition, we will interview representatives of MINEDUC and FHI360 (the implementer). Activity 3: We will interview key informants at the national level, in two participating regions, and possibly in subregions, depending on the decision-making locus.

Sampling Procedure

Activity 1, quantitative impact evaluation

We will estimate the causal impacts of the Éxito Escolar activity using a randomized controlled trial. We have randomly assigned school districts to a treatment group or control group. We conducted random assignment via a public lottery in Guatemala.

Ahead of the public lottery, we formed strata to permit stratified random assignment. We first stratified districts by department and then by 2013 standardized scores in language and communications and in mathematics to ensure balance on those key characteristics and to more precisely estimate the impacts on learning. To group schools with similar test scores, we used district-level means (averaged across language and communication and across math) and sorted schools within departments by their 2013 test scores. For schools with no 2013 test scores, we imputed their scores using their municipality's mean score. The schools in our sample have test scores that are below the national average. The district means, averaged across language and communication and math, are 1.11 to 0.07 standard deviations below the national mean.

If a disproportionate number of large districts (districts with many teachers) ended up in the control group, we might have less than the implementer's minimum of 80 percent of teachers in the treatment group. To avoid this, we reassigned some districts to different strata. These reassignments involved clustering the largest districts in strata with other large districts. This way, if a large district is assigned to the control group, other large districts in the same stratum are guaranteed to be assigned to the treatment group.

We selected replacement schools for MINEDUC and Espirálica to visit if they are unable to collect data in one of the selected schools. Espirálica staff will select replacement teachers randomly if the teacher they intended to interview is unavailable.

We used power calculations to estimate the minimum detectable effect size (MDES) that we would be able to identify with our design. We calculated intracluster correlation and variance reduction using 2009 and 2013 test-score data from students in public lower-secondary schools in the five study departments. To calculate the MDES, we assumed 80 percent power, a two-tailed test, and a 5 percent significance level. For mathematics, we estimated that the covariates we plan to include in

the impact regression model will explain 4 percent of the individual-level variance in the outcome variable and 73 percent of the group-level variance in the outcome variable. The intracluster correlation coefficient for students in the districts is 0.08. For language and communications, we estimated that the covariates will explain 6 percent of the individual-level variance in the outcome variable and 84 percent of the group-level variance and that the intracluster correlation coefficient for students in the districts is 0.19.

The results pertain to students from one grade. We assume that we will have test-score data for 75 percent of students in sample schools.

We will use two methods to estimate impacts on students' test scores: intent to treat (ITT) estimates, which estimate the impact of being assigned to the treatment group, regardless of participation in treatment, and local average treatment effect estimates, which are inflated to adjust for schools in the treatment group that fail to receive treatment and schools in the control group who do receive treatment. MDES are lower for ITT estimates than for LATE estimates. For ITT estimates, we estimated an MDES of 0.12 for math and 0.13 for language and communication. For LATE estimates, we assume 70 percent of students in treatment schools receive their intended treatment, but 10 percent of students in control schools also receive treatment. Our MDES for LATE estimates increases to 0.20 for math and 0.22 for language and communication. For teachers' time on task, we estimated that we will be able to detect an increase in the percentage of time on task of 10 to 11 percent with the ITT estimates or of 17 to 18 percent with the LATE estimates.

Activity 1, qualitative evaluation of implementation

We will use purposeful, maximum variation sampling, in which we will select participants based on their affiliation with a high-, medium-, or low-performing school. We will use baseline test-score data to identify schools' performance levels at baseline and use midline test-score data to identify performance levels at endline. This approach will enable us to identify themes that occur consistently across schools and participants, regardless of variation in performance. It will also enable us to identify issues that could be unique to any one group, such as high-performing schools or treatment schools (Patton 1990). Mathematica will rely on MINEDUC administrative data (including test scores) to identify thresholds for high-, middle-, and low-performing schools.

Deviations from Sample Design

We have no deviations from the sample design to report at this time.

Response Rate

We will update this field once baseline data collection is complete

Weighting

Activity 1:

Weighting: Because we drew a simple random sample of schools and are including all students from selected schools in test-score data collection, we will not need to use survey weights to analyze the test-score data. We will evaluate the need for response weights.
For analyses of teacher data at the school level, we will have no need for survey weights. But for analyses of teacher data at the teacher level, we will use survey weights to account for the fact that teachers at smaller schools are more likely to be selected than teachers at larger schools.

Activity 3:

We do not anticipate a need to use weighting in our analysis.

Questionnaires

Overview

1. Impact and implementation evaluations of Activity 1

We will use diverse methods to gather the data required for the impact and implementation evaluations: primary survey and qualitative data, MINEDUC's administrative and student assessment data, and FHI360's monitoring and evaluation data.
Primary quantitative data collection: surveys and classroom observations
At baseline and endline, we will administer teacher and school director surveys and conduct Stallings classroom observations in the classrooms of teachers selected for the survey. For both rounds of data collection, we will gather data in the 332 schools selected for the data collection sample. Our data collection firm, Espirálica, piloted the survey and classroom observation instruments in April with 25 schools in departments located near the study departments. We worked with Espirálica to revise the instruments based on the results of the pilot. Espirálica will use electronic data collection to reduce data-entry errors and to accelerate the submission of the data. Mathematica will review data as they are collected to provide rapid feedback on any problems that arise.
Content of teacher and school director surveys:

The baseline and endline surveys will be administered in Spanish and will include the following modules:

- Demographic characteristics and qualifications. School directors and teachers will report on their gender, age, and ethnic group; languages spoken; level of education; and years of experience in the education field.
- Professional development. We will ask school directors and teachers about their recent professional development training on topics covered by Éxito Escolar, such as language and communications, mathematics, natural sciences, leadership, and administration.
- Work conditions. We will ask school directors and teachers about their workload, including the number of hours spent on administrative, teaching, and professional development activities. Teachers will report on the grades and subjects they teach and the number of students per classroom. Work conditions are likely to affect school directors' and teachers' willingness and ability to complete the training program and use the new pedagogic techniques.
- Interest and self-efficacy. At baseline, we will ask school directors and teachers about their interest in participating in Éxito Escolar's subactivities and their perceived self-efficacy for completing them.
- Pedagogic practices. Teachers will describe their classroom practices and beliefs about best practices in teaching.
- Pedagogic support. School directors and teachers will report on the level of pedagogic support received in the past 12 months.

Primary qualitative data collection: focus groups and interviews in English and Spanish:

The implementation study will draw on two rounds of qualitative data collected through (1) interviews with key stakeholders (MCC, MINEDUC officials, implementers of professional development programs for teachers, and school directors) and (2) focus groups with project participants, including teachers, students, and parents. We will develop protocols for qualitative data collection that are customized to each data collection method (interview or focus group) and respondent, but the protocols will all cover similar topics related to the research questions, such as demographics and educational background, attitudes about professional development programs for teachers, facilitators of and barriers to implementation, pedagogic support, perceptions of training and capacity building (teacher diagnostics), the role of parent councils, implementation of assessments, and perceptions of school networks. Before the first round of data collection, Espirálica, under Mathematica's oversight, will pilot test the protocols for sequencing, logic, and comprehension.

Administrative data:

We will use data collected by MINEDUC's planning office, DIPLAN, to measure key school characteristics and student enrollment data. Important school characteristics include schools' modality (such as national basic education institutes, cooperative, and telesecondary); sector (public or cooperative); and student enrollment data (enrollment, dropout, grade repetition, transition from primary to lower- secondary school, and transition from lower- to upper-secondary school).
To assess the quality of the administrative data, Mathematica will apply internal data quality review processes to all administrative data sets we use. We will obtain copies of the data collection forms to improve our understanding of the data sets provided and to check the data for consistency, including looking for outliers and missing values. We will also ask MINEDUC to verify any suspicious data points. We may not be able to correct inaccurate data, but we will try to identify inaccuracies and will describe any consequent limitations when we report our results.

Monitoring and evaluation data:

We will use FHI360's monitoring and evaluation data to deepen our understanding of the implementation of Activity 1. Key monitoring and evaluation data that we expect to use include the number of teachers and school directors who express interest in participating in Éxito Escolar activities, the number who enroll, participation rates in the activities, and the implementation of the activities. We also plan to use monitoring and evaluation data on the outcomes of the activities, such as the formation of educational networks, parent organizations, and learning communities.

2. Performance evaluation mixed-methods approach for Activity 3 (Strengthening Institutional and Planning Capacity activity)

The first component of the evaluation of Activity 3 is a trend analysis, for which we will use secondary data to examine changes over time in key outcomes related to the project (for example, changes in budget allocations to secondary education and changes in teacher hiring and retention). We will draw on data from all departments in Guatemala (Activity 3 is national in scope) and will rely on project-monitoring data from the MCC monitoring and evaluation plan.

The second component is qualitative analysis, for which we will draw on three rounds of key informant interviews with stakeholders and two rounds of focus groups with program participants such as teachers, school directors, and university faculty. The interviewees will be national staff, departmental staff, and project implementers. The focus group participants will be teachers, school directors, and parents from the five departments where Activity 1 will be implemented. The first round of data collection (in the first two quarters of 2018) will capture information about the current status of the education system (from an institutional perspective) and perceptions of the teacher hiring and recruitment strategy, budget allocations for secondary education, and the needs of the secondary education system. The second round of data collection (in the first two quarters of 2019) will capture mid-term trends and perceptions from stakeholders to help guide the trend analysis. The final data collection will take place at the end of 2020 alongside data collection for Éxito Escolar.

Data Collection

Data Collection Dates

Start	End	Cycle
2018-05-21	2018-06-29	Activity 1, baseline quantitative data collection
2018-10	2018-10	Activity 1, baseline qualitative data collection
2018-07	2018-07	Activity 3, baseline data collection
2019-06	2019-06	Activity 3, first follow up data collection
2020-10	2020-10	Activity 1 and 3, endline quantitative and qualitative data collection

Data Collection Notes

Baseline quantitative data collection is ongoing and is expected to be completed by the end of June.

Our data collection firm, Espirálica piloted the survey and classroom observation instruments in April with 25 schools in departments located near the study departments. We worked with Espirálica to revise the instruments based on the results of the pilot. Espirálica will use electronic data collection to reduce data-entry errors and to accelerate the submission of the data. Mathematica will review data as they are collected to provide rapid feedback on any problems that arise.

All enumerators received a full week of training on the surveys and, when applicable, a full week of training on administering the classroom observation instrument. Espirálica evaluated enumerators at the end of training and selected those who met its selection criteria to engage in field work.

Questionnaires

1. Impact and implementation evaluations of Activity 1

We will use diverse methods to gather the data required for the impact and implementation evaluations: primary survey and qualitative data, MINEDUC's administrative and student assessment data, and FHI360's monitoring and evaluation data.
Primary quantitative data collection: surveys and classroom observations
At baseline and endline, we will administer teacher and school director surveys and conduct Stallings classroom observations in the classrooms of teachers selected for the survey. For both rounds of data collection, we will gather data in the 332 schools selected for the data collection sample. Our data collection firm, Espirálica, piloted the survey and classroom observation instruments in April with 25 schools in departments located near the study departments. We worked with Espirálica to revise the instruments based on the results of the pilot. Espirálica will use electronic data collection to reduce data-entry errors and to accelerate the submission of the data. Mathematica will review data as they are collected to provide rapid feedback on any problems that arise.
Content of teacher and school director surveys:

The baseline and endline surveys will be administered in Spanish and will include the following modules:

- Demographic characteristics and qualifications. School directors and teachers will report on their gender, age, and ethnic group; languages spoken; level of education; and years of experience in the education field.
- Professional development. We will ask school directors and teachers about their recent professional development training on topics covered by Éxito Escolar, such as language and communications, mathematics, natural sciences, leadership, and administration.
- Work conditions. We will ask school directors and teachers about their workload, including the number of hours spent on administrative, teaching, and professional development activities. Teachers will report on the grades and subjects they teach and the number of students per classroom. Work conditions are likely to affect school directors' and teachers' willingness and ability to complete the training program and use the new pedagogic techniques.
- Interest and self-efficacy. At baseline, we will ask school directors and teachers about their interest in participating in Éxito Escolar's subactivities and their perceived self-efficacy for completing them.
- Pedagogic practices. Teachers will describe their classroom practices and beliefs about best practices in teaching.
- Pedagogic support. School directors and teachers will report on the level of pedagogic support received in the past 12 months.

Primary qualitative data collection: focus groups and interviews in English and Spanish:

The implementation study will draw on two rounds of qualitative data collected through (1) interviews with key stakeholders

(MCC, MINEDUC officials, implementers of professional development programs for teachers, and school directors) and (2) focus groups with project participants, including teachers, students, and parents. We will develop protocols for qualitative data collection that are customized to each data collection method (interview or focus group) and respondent, but the protocols will all cover similar topics related to the research questions, such as demographics and educational background, attitudes about professional development programs for teachers, facilitators of and barriers to implementation, pedagogic support, perceptions of training and capacity building (teacher diagnostics), the role of parent councils, implementation of assessments, and perceptions of school networks. Before the first round of data collection, Espirálica, under Mathematica's oversight, will pilot test the protocols for sequencing, logic, and comprehension.

Administrative data:

We will use data collected by MINEDUC's planning office, DIPLAN, to measure key school characteristics and student enrollment data. Important school characteristics include schools' modality (such as national basic education institutes, cooperative, and telesecondary); sector (public or cooperative); and student enrollment data (enrollment, dropout, grade repetition, transition from primary to lower- secondary school, and transition from lower- to upper-secondary school).
To assess the quality of the administrative data, Mathematica will apply internal data quality review processes to all administrative data sets we use. We will obtain copies of the data collection forms to improve our understanding of the data sets provided and to check the data for consistency, including looking for outliers and missing values. We will also ask MINEDUC to verify any suspicious data points. We may not be able to correct inaccurate data, but we will try to identify inaccuracies and will describe any consequent limitations when we report our results.

Monitoring and evaluation data:

We will use FHI360's monitoring and evaluation data to deepen our understanding of the implementation of Activity 1. Key monitoring and evaluation data that we expect to use include the number of teachers and school directors who express interest in participating in Éxito Escolar activities, the number who enroll, participation rates in the activities, and the implementation of the activities. We also plan to use monitoring and evaluation data on the outcomes of the activities, such as the formation of educational networks, parent organizations, and learning communities.

2. Performance evaluation mixed-methods approach for Activity 3 (Strengthening Institutional and Planning Capacity activity)

The first component of the evaluation of Activity 3 is a trend analysis, for which we will use secondary data to examine changes over time in key outcomes related to the project (for example, changes in budget allocations to secondary education and changes in teacher hiring and retention). We will draw on data from all departments in Guatemala (Activity 3 is national in scope) and will rely on project-monitoring data from the MCC monitoring and evaluation plan.

The second component is qualitative analysis, for which we will draw on three rounds of key informant interviews with stakeholders and two rounds of focus groups with program participants such as teachers, school directors, and university faculty. The interviewees will be national staff, departmental staff, and project implementers. The focus group participants will be teachers, school directors, and parents from the five departments where Activity 1 will be implemented. The first round of data collection (in the first two quarters of 2018) will capture information about the current status of the education system (from an institutional perspective) and perceptions of the teacher hiring and recruitment strategy, budget allocations for secondary education, and the needs of the secondary education system. The second round of data collection (in the first two quarters of 2019) will capture mid-term trends and perceptions from stakeholders to help guide the trend analysis. The final data collection will take place at the end of 2020 alongside data collection for Éxito Escolar.

Data Collectors

Name	Abbreviation	Affiliation
Espirálica Research & Consulting		

Supervision

Activity 1:
The quantitative data collection for Activity 1 started on May 21, 2018.

Espirálica Research and Consulting is conducting the primary data collection, including surveys, classroom observation, focus groups, and some interviews (Mathematica staff will conduct remaining interviews). Cesar Valenzuela is the president of Espirálica. Daniela Cruz is the field director. Mathematica staff will supervise Espirálica's work. The data collection team consists of 2 teams (27 members). Each team is comprised of 1 supervisor, 8-9 enumerators and observers and 3 drivers.

The supervisors are reporting to the Country Coordinator, Field Coordinator and the Survey Director.

The supervisor is performing the following types of tasks:
· Oversee data collection in the field, including interacting with local community leadership and school officials to facilitate data collection, assuring proper dispatching of data collection personnel to the correct study sites, determining which data collection personnel will be assigned specific cases or tasks, and ensuring that those cases or tasks are completed
· Conduct quality assurance checks on each completed survey and observation as soon after the task is completed as possible, ideally while data collection personnel are still in the field and could return to respondents if errors so require
· Observe the survey and observation data collection and correct any data collection problems, including retraining field staff when necessary for systematic problems or changes to protocols or instruments
· Inform the field coordinator or survey director immediately about any problems encountered
· Perform other tasks and assume other responsibilities as requested by Mathematica

Mathematica is supervising Espirálica's data collection work.

DIGEMOCA (Ministry of Education, Guatemala) will gather student test score data.

FHI360 will gather monitoring and evaluation data.

Data Processing

Data Editing

Activity 1:

Espirálica and Mathematica will take several steps to clean the survey and classroom observation data. In the field, data collection supervisors review all completed surveys and classroom observations every day to verify that the data are complete and free from errors. If supervisors find missing data or errors, enumerators or observers follow up with respondents as necessary to attempt to resolve the issue. In addition, for 10 percent of the surveys, Espirálica calls respondents to verify responses to a set of key questions, providing a second opportunity to detect errors in the data. Once Mathematica staff receive the data, they will also check the data for duplicate or missing observations and review all variables to verify that data are complete, that response values are reasonable, and that skip patterns are followed, following up with Espirálica to resolve any issues. Furthermore, Mathematica will take additional steps to make the data conducive to analysis, including coding frequently listed responses to “other” responses.

Mathematica will use data collected by others for both the impact and implementation evaluations. Secondary data sources include MINEDUC's student test-score data and administrative enrollment data and FHI360's monitoring and evaluation data and the results of the teacher knowledge test they will administer. We will review all data sources for missing or inconsistent data. We will also determine what data quality checks are appropriate when we see the data, but we anticipate conducting range checks and reviewing the distribution of the data to check for inconsistencies. We will follow up with the organizations that collected the data to attempt to resolve any issues found.

Espirálica will produce transcriptions of focus groups and interviews. Mathematica will compare a selection of the transcripts against audio recordings to confirm that the transcriptions are complete and accurate.

Activity 3:

We will not collect primary quantitative data for Activity 3. However, we will review the quantitative administrative data we receive for potential missing data or inconsistencies. We will contact the producer of the data to investigate any issues found and identify potential solutions.

As in the process described for Activity 1, Espirálica will produce transcripts of interviews conducted for Activity 3. We will compare a selection of the transcripts against audio recordings to confirm that the transcriptions are complete and accurate.

Other Processing

Activity 1:

All survey and observation data is collected electronically, via tablets. The data collection firm is reviewing data for consistency as it is collected electronically from the field, cleaning the data (as necessary) and transmitting the data electronically to Mathematica for review as it is available. The data collection firm will send the data electronically to Mathematica for data checks after the first 5 percent of cases have been collected. Mathematica will conduct data verification checks to ensure that the data collected and entered are reliable and accurate. If inconsistencies are found, the data collection firm will be responsible for implementing the necessary corrections.

Data Appraisal

Estimates of Sampling Error

Activity 1:

We will use regression analysis to estimate the impacts of Activity 1 on outcomes of interest. We will report the standard errors of our estimates as well as the p-values associated with our estimates. In addition to the treatment indicator, our regression equations will include baseline covariates to control for chance differences between treatment groups at baseline, and strata fixed effects to account for the stratified random assignment. We will cluster standard errors at the school district level to account for the clustered randomization.